VOLUME 64
NUMBER 10

WHOLE NO. 316
1950

Psychological Monographs:

General and Applied

Combining the *Applied Psychology Monographs* and the *Archives of Psychology* with the *Psychological Monographs*

HERBERT S. CONRAD, *Editor*

Experimentally Induced Variations in Rorschach Performance

By
EDITH LORD
Mental Health Division
Arizona State Department of Health

Accepted for publication, April 23, 1950

Price $1.00

Published by
THE AMERICAN PSYCHOLOGICAL ASSOCIATION
1515 MASSACHUSETTS AVE., N.W., WASHINGTON 5, D.C.

ACKNOWLEDGMENTS

ACKNOWLEDGMENT is made to the following persons for valuable suggestions and criticisms of this study: J. P. Guilford, William Grings, Bruno Klopfer, Louis P. Thorpe, Georgene Seward, S. M. Wesley, George R. Bach, Saul Rosenzweig, and Herbert S. Conrad.

CONTENTS

I. INTRODUCTION

DURING the past decade many new personality tests have been devised, little-used old tests have been revised, popular tests have been polished, refined, and expanded. College catalogues throughout the United States have lengthened lists of courses offered in psychology to include instruction in the administration and interpretation of personality tests in general, of "projective techniques" in particular, and of the Rorschach Ink Blot Test specifically.

Literally hundreds of research reports and observational articles have been published on the Rorschach and its clinical use. The bulk of this literature has been concerned with Rorschach patterns as they relate to psychiatric syndromes, to social or cultural groups, to successful and unsuccessful students, workers, or military groups. The practical or empirical success of the instrument has been widely reported. The operational results of the Rorschach as a diagnostic instrument have been repeatedly acclaimed.

There is, however, a dearth of laboratory investigation of this important and widely-used clinical tool. In fact, there is even a question as to whether the test actually is a projective technique. Schachtel (31) offers an excellent critique of the projection hypothesis of the Rorschach test, and Bellak (3) states flatly that Rorschach's test is *not* a projective method. More clarification of these philosophical and theoretical areas underlying the Rorschach is definitely indicated.

Too, the paucity of statistical studies of reliability and validity, or of adequate norms, plus the existence of contradictory reports on these crucial points has given rise to a diapason of criticism. This cacophony has been particularly painful to the ears of the conscientious clinical psychologist who daily uses the Rorschach for diagnosis, for recommendations for therapy, for vocational or educational guidance and counseling, etc.

As Hsü (16) points out, there are two popular and disparate attitudes toward the Rorschach held by equally psychologically sophisticated persons. One maintains that the prime importance of the instrument is its revelation in symbolic terms of the personality of the individual, considered as a unique universe; the other stoutly maintains that this test, like all other psychological measuring devices, should meet at least minimum standards of statistical reliability and validity.

Hunt (17) criticizes the over-emphasis in clinical psychology on quantitative data rather than qualitative, and bases his hopes for the future progress of clinical testing on an increased attention to qualitative behavior of the person tested rather than on quantitative results of the test.

Qualitatively, does a subject, in fact, satisfy one of the basic assumptions underlying all projective techniques—that his concepts reveal his consistent way of organizing experience, as measured by the ideas and feelings he projects into "meaning-free" or ambiguous stimuli? Is a subject reacting only to the "unstructured" or "semi-structured" ink blots on the Rorschach test, or is he reacting to the total perceptual and social field, which includes the administrator of the test and the affective tone of the administrative situation? If these latter two variables are altered, will the subject's perception of the ink blots vary significantly? Can a subject conceal his way of

organizing experience or simulate a pattern not his own? Can an administrator induce an altered Rorschach pattern by explicit or implicit suggestion?

There have been some very interesting explorations of these last two questions. Fosberg (11) studied attempts to fake Rorschach results, and also (12) examined protocols obtained under varied instructions. He demonstrated that psychologically sophisticated subjects, given a neutral or control Rorschach, could not conceal their basic test patterns on two subsequent tests in which they tried to give the best, then the worst, possible impression.

Schachtel (30) emphasizes the importance of a subject's personal definition of the test situation and its demands. He also strongly presents his personal conviction that the three most common subjective definitions of the Rorschach situation are as authoritarian, competitive, or resistance situations. These definitions, he maintains, influence Rorschach results.

Levine (22, 23) reports a study of a series of protocols produced by the same subject while experiencing different situations hypnotically induced. There were marked changes in the Rorschach records of the various situations, each taking on the coloring considered characteristic of a person experiencing the situation created by hpnotic suggestion.

Luchins (27) studied the influence on Rorschach responses of situational and attitudinal factors. Among his findings was the marked need for the development of methods which would prepare a subject for the Rorschach test. A significant factor in the Rorschach results he analyzed proved to be misunderstandings of directions and subjects' hazy concepts of what was expected of them in the testing situation.

Wilkins and Adams (33) studied results of Rorschachs obtained under hypnosis, under sodium amytal, and in the normal, waking state. They found that both the chemically and psychologically induced abnormal states contributed to increased productivity on the Rorschach among subjects described as "overly cautious" or "fearful."

An interesting variation in the sort of study conducted by Levine is that of Bergman, Graham, and Leavitt (4) who reported the Rorschach records of a hypnotically regressed patient, diagnosed as a conversion hysteric. They found that the protocols varied with the level of hypnotic regression but closely followed the expectancies growing out of other clinical data.

Cofer (9) studying the changes in responses given by a patient with organic involvement while under hyoscine, found that some of the cards are more sensitive than others to situational variations.

A recently published research project of the Army Air Forces (32) included an analysis of the influence of the examiner on the number of responses obtained on Rorschach records. Of 36 t-ratios between the means of nine examiners, 12 were significant at the 1 per cent level of confidence and three at the 5 per cent level. Clearly this one scoring symbol (R)[1] would seem to vary with the examiner.

Keeping this last finding in mind, and hypothesizing a continuum between the findings of Levine and Fosberg, one wonders whether the non-hypnotized subject ever experiences attitudes toward the administrator or mood reactions to the total administrative situation which are adequate to alter the determinants of his Rorschach record significantly. This, in brief, is the problem which led to the development of the present study.

---

[1] An explanation of the Rorschach scoring symbols used in this study is given in the Appendix.

II. THE PROBLEM

STATEMENT OF THE PROBLEM. This study of the influence of negative and positive rapport conditions on Rorschach performance is based on the hypothesis that the perception of ambiguous or "unstructured" stimuli is not influenced by the perceiver's affective reaction toward the total stimulus field, which includes the administrator—the person presenting the "unstructured" stimuli. If an experimental test proves this hypothesis false, implications are legion for the interpretation of test results and for the training of administrators.

Of importance is the possibility that this study—a test of the foregoing hypothesis—may reveal whether or not the administrator is a significant part of the stimulus field during the administration of projective tests.[1] The theory behind projective testing is that the subject structures stimuli presented to him in non-structured or semi-structured form. If a subject is projecting meanings, ideas, and feelings into ambiguous stimuli uninfluenced by the person presenting the stimuli or by the affective tone of the administrative situation, there should be no significant variation in his structuring of the stimuli if either the administrator or the affective tone of the administration is varied.

The problem of transference arises in all therapies, regardless of type, and the findings of this project might conceivably throw some needed light on the subject of transference-elements. If the requirements for positive transference could be determined or the situations in which it flourishes could be measured, the selection and training of therapists could be more clearly defined. It is thought that the present study may hold implications for this related problem.

Outside the realm of clinical psychology, this study may have implications for social psychology. Too, the manner in which inter-personal relationships affect perceptual processes is of great current interest to systematic psychology (14) as well as to the field of clinical psychology. While the present study does not answer the questions here posed, it does provide implications for future research in these areas.

Within the framework of the therapeutic situation, there may be at work individual variables which cause one subject to function more "normally" in a threatening environment, another to function more productively or adequately in an environment of heightened acceptance or permissiveness. In so far as the therapist's personality may structure these mood-situations, the results of this project could be revealing.

Designing an experiment to test whether Rorschach records will vary with experimental variation of administrator and tone of administration, one must ask himself, from a theoretical point of view, what variations might be expected. Remembering the experiments of Levine, one might hypothesize radical shifts in protocols from one administration to another if the administrations were adequately loaded with different affective tones. Remembering the experiments of

[1] Cognizance is taken here of the fact that one cannot amplify specific findings related to the Rorschach ink blots into generalizations concerning projective techniques as a category; however, it is thought that any light on the variability or stability of one projective technique may, at least, suggest areas of desirable investigation for all other clinical tools falling within the "projective" category.

Fosberg, wherein psychologically sophisticated subjects proved incapable of successfully simulating the Rorschach records of personality-types other than their own, one might predict no significant variation in total protocol, regardless of administrator or of type of administration.

Somewhere between these extremes, one may ask what sort of variation might be expected (a) if a group of subjects were given control Rorschach administrations, later were given positively loaded[2] Rorschach administrations, and still later negatively loaded administrations; (b) if a group of subjects were given Rorschachs by three different administrators at intervals of four to six weeks. Following are the hypotheses which seem to follow from the cited relevant studies, and which seem reasonable from the standpoint of personality dynamics and the Rorschach symbols or projective response-categories which are widely believed to be related to these dynamics.

1. There will be a variation in the number of responses elicited by the various administrators, regardless of affective variable or sequence of administration. This hypothesis is based on the A.A.F. study, cited above.

2. There will be no consistent variation in number of responses between the first, second, and third administrations; i.e., no increase or decrease in number of responses with successive administrations. However, the variation in number of responses will be related to a subject's initial performance. If a subject gives thirty or fewer responses on his first test, he will increase the number with subsequent testings; if he gives seventy or more on the first, he will decrease his responses on subsequent testing.[3]

3. Even among "normal" subjects, the alterations attendant upon variations of affective tone of administration will be largely dependent upon individual differences, the more stable personality records showing less change, the more labile records showing more change. It is predicted that approximately a third of the subjects' records will show no differences or only minimal changes regardless of administrator, order of administration, or experimentally varied affective tone of administration; a third will be extremely variable, fluctuating with administration order, administrator, and tone of administration. This stability or lability will be reflected in the constancy or shift of the M : sumC ratio.[4]

4. Regardless of the order of administration or the affective loading, there will be some variation in the records obtained by the three administrators which will be a function of the different personalities of the administrators. Since this factor is an unknown variable, it is not possible to predict the anticipated variations specifically. This hypothesis is a logical outgrowth of the known variation of the number of responses with examiner differences (32). Responses are scored as to location areas and determinants. An increase in responses insures an increase in one or another of these scoring categories. Furthermore, since number of responses has been shown to vary with examiner differences, it is not

[2] The word *loaded* is here used to mean *weighted* or *slanted*. See IV, C for complete exposition of this term as herein employed.

[3] Since thirty to seventy responses is usually considered the normal range, this hypothesis is based on an anticipation of the operation of central tendency.

[4] The particular choice of thirds in this hypothesis is based on verbal prediction made by Dr. Bruno Klopfer.

unreasonable to hypothesize that other scoring categories of the Rorschach may be equally sensitive to administrator differences.

5. Regardless of the administrator or the sequence of administration, there will be variations in the protocols produced under the three different administrative situations: neutral or control, positive affective loading, negative affective loading. Specific variations anticipated are the following:

a. Negatively loaded administration.

(1) Greater scattering of response determinants among the signs k, K, Fc, and C', indicating a withdrawal from the painful situation (k or K), a depression of a situational sort (C'), and a sensitive approach to the total stimulus environment (Fc) as a way of handling anxiety provoked by the situation. These signs will tend to increase at the cost of the movement and color responses. Fm will appear with greater frequency as an indicator of tension in the face of a threatening situation.

(2) There will be an increase in white-space responses (S) as an indication of negativistic or oppositional tendencies.

(3) There will be an increase in F% and in FK + F + Fc/R as a reflection of an increased effort to master the situation through intellectual control or through this control supported by heightened sensitivity and introspection.

b. Positively loaded administration.

(1) There will be a greater emotional expression, or emotional relationship with the environment in this more permissive situation; this will be reflected in an increase in sumC.

(2) There will be an increase in movement responses, M and FM, indicating a wider play of creativity and a greater freedom of expression of "Id-drives."

III. PROCEDURE

AN EXPERIMENTAL design was constructed which aimed toward controlling as many variables as possible. The basic unit of the study was 36 males between the ages of 19 and 27. They were selected at random from a list of approximately 200 college sophomores enrolled in an introductory course in general psychology. Each subject was given three Rorschachs at four to six weeks' intervals. Each of the three Rorschachs was administered by a different person.

The three administrators, A, B, and C, were females, not different in any grossly apparent way; i.e., no one of the three was outstandingly fat or thin, ugly or beautiful. Administrators A and C were brunettes, B was a reddish-blonde. The age range among the three was twelve years, with B the youngest and C the oldest. Each administrator holds a degree in psychology and has had several years' experience in the use of the Rorschach according to the Klopfer technique.

The Rorschach tests were given in a systematic or rotated order so that twelve subjects received control tests first, twelve second, and twelve third. Twelve received negative tests and twelve received positive tests first. The second and third testing sessions were similarly varied.

No subject was tested twice by the same person. Twelve were tested first by Administrator A, second by B, and third by C. The other administrations were distributed similarly. Each administrator used exactly the same technique for the administration of the Rorschach Test, regardless of the affective loading of the pre-test situation and regardless of whether the administration was the first, second, or third for the subject.

The negative and positive situations included two simple card-sorting tests administered prior to the Rorschach test for the purpose of setting the affective tone of the session. Identical directions, terminology, and techniques were used in both the negative and positive pre-tests; however, the tone of the administrations was differently slanted. Specific directions for the administration of the three sessions were followed explicitly and without variation by the three administrators. These directions were as follows:

ADMINISTRATION

Neutral. Standard Rorschach procedure will be used. The administrator will be courteous but business-like in manner. She must attempt to avoid either negative or positive affective loading of the situation. There will be no pre-tests nor any gathering of biographical data.

Negative administration. The administrator will assume the role of a harsh, rejecting, authoritarian figure. She must be deliberately unconcerned about the subject, not look at him while asking questions, preparing tests, or giving directions; never smile, give directions in a voice of dictatorial harshness, make every "Hm!" sound like a sneer.

Positive administration. The administrator will be personally warm, charming, appreciative in manner. She must look at subject with a smile while asking questions, preparing tests, or giving directions in an encouraging tone of voice, making every "Hm!" sound like a compliment for work well done.

The following administration will be used verbatim for *both* negative and positive administrations:

Are you Mr. Blank? Come with me! This room. Sit here.

What is your full name? You have an address? Freshman, Sophomore, Junior? Your birthdate?

Are you familiar with an ordinary deck of playing cards? (Place deck face down, with one black nine on the bottom of the pack.) When I say GO turn the cards, one at a time, and pull out all the black nines and the red tens. Put the black nines here and the red tens here (indicate areas with gestures). Work as fast as you can. Ready? Go! (Answer any questions only by repeating words in the foregoing instructions. Note any cheating, i.e., number of times subject turns two or more cards at a time; also note any errors in carrying out instructions or any requests for repetition of instructions.) Record time.

Hm! Some speed! (Leave nines and tens; shuffle remaining cards; replace, face down, before subject.)

Now, when I say GO, turn the cards, one at a time, and see how many face cards—Jacks, Queens, and Kings—you can pull out before I say STOP. Put the red face cards on the black nines and the black face cards on the red tens. Ready? Go! (Allow thirty seconds.)

STOP! (Count face cards and record number.) Hm! Not bad!

(The Rorschach Test will follow using the same instructions as for the neutral administration.)

As the experimental design reveals, each subject follows a unique pattern. Consequently, if a subject were lost at any stage of the experiment, a substitute would have to be started from the very beginning. In order to insure completion of the design, the entire pattern was started in duplicate; i.e., seventy-two subjects were given first Rorschachs. Despite this precaution, there were several cases wherein both of the pair were lost, and a third subject (with his parallel substitute) was started through the design. The principal cause for loss appeared to be the subject's reactions to the negatively administered session. Following this session, the subjects seemed more difficult to contact for testing appointments and frequently broke appointments or failed to appear as scheduled. Thirteen of the original seventy-two subjects were dropped following the negatively administered test because they had failed to keep from three to five appointments for the subsequent test. Two subjects were dropped for the same reason following positively administered sessions. Eight of the thirteen were given negatively loaded tests during the first session, five during the second. The losses, by examiner, were approximately equal, two losing four subjects each, the third losing five after negative administrations. The only reason for dropping these subjects, rather than persisting until an appointment was kept, was the time factor. If a period longer than six weeks elapsed between sessions, the subject was dropped and his parallel substituted in the pattern.

With this evidence in mind, one may speculate that the differences between negative results and positive or neutral results might have been greater if methodology had permitted retention of the negative records of the subjects who apparently rejected the entire project after experiencing the negative administration.

A few subjects were lost for other reasons. Three moved from the community before completing the three sessions; two were ill between sessions and unable to meet the requirement of a maximum time-lapse of six weeks between tests.

In every other case, testing of parallel partners was dropped as soon as the principal subject had completed his pattern and thereby precluded the need for a substitute.

As Rorschach records were accumulated, they were coded and mixed. The one hundred eight records were then scored without reference to subject, name of administrator, type or order of administration. Only main responses and determinants were used in the analyses of the records for two reasons: (1) incorporation of additional responses and determinants would unprofitably overcomplicate the study; (2) there are marked differences among exponents of the Rorschach method (e.g., Beck, Buhler, and Klopfer) as to the most desirable use to make of these additional determinants in scoring and in interpretation.

IV. RESULTS OF THE EXPERIMENT

In addition to the loss of subjects following negative administrations, there is other evidence that the negative situation was actually traumatic to the subjects. Although the very simple card-sorting pre-tests were designed merely to set the administrative tone of the administration for the Rorschach, behavior on the pre-tests suggests that in many cases the subjects were reacting differently by the time the pre-tests were introduced. The initial greeting, without a smile or a direct look, plus the manner and tone of voice in which the four personal-data questions were asked, apparently caused the subjects some concern.

A. Behavior on Pre-Tests

Not a single subject on the positively-loaded administrations misunderstood the pre-test directions or made an error in carrying them out. Twelve of the thirty-six subjects made from one to six errors on the negative administrations, with a total of twenty-five errors. Failure to understand directions was counted as an error if the subject requested repetition or clarification. If he demonstrated his confusion through incorrectly carrying out the directions, he was not corrected, but the error was noted.

During the Rorschach administration, the writer recorded every audible sigh or audible laugh. The laughs and sighs elicited by this one administrator give further qualitative evidence that the subjects were, in fact, reacting differently to the different affective loadings of the administrations: Eleven of the thirty-six subjects sighed aloud from one to seven times (total of thirty-two sighs) during the free-association period of the Rorschach, i.e., during the initial presentation of the plates. One subject sighed aloud during the positive, two during the neutral, and eight during the negative administrations. Seven subjects laughed aloud once or twice with a total of nine audible laughs; three laughed during the positive, two during the negative, and two during the neutral administrations. No effort was made to tabulate smiles and merely visible sighs.

One observes that audible sighs occur with greater frequency during a cold or rejecting situation than during a warm, permissive situation or a neutral situation, but that laughing aloud occurs with approximately equal frequency during the three different situations.

B. Variation in Number of Responses

The hypothesis concerning successive increase or decrease of number of responses as a function of the number given on the initial test held well for two cases which initially gave more than seventy responses but failed to hold consistently for twenty-six cases giving fewer than thirty responses at the outset.

There is an apparent bias in testing response limits exceeded by such an uneven number of subjects as exceeded our limits; i.e., two versus twenty-six. The explanation of this procedure rests in the earlier hypothesis based on erroneous original expectancy as to number of responses. Rorschach literature quite generally states that thirty responses can be considered the minimum expectancy with normal, intelligent, adult subjects, seventy responses the maximum expectancy. The failure of our subjects to conform to this expectancy is discussed later under the subheading, "Number of responses."

C. Erlebnistyp

Hermann Rorschach's theory of personality rests on the foundation of what he called *Erlebnistyp,* a word which has been variously translated but which is consistently rendered *Experience Type* in the English translation of Rorschach's original paper. He (29) makes reference to the problem, so labeled, throughout his *Psychodiagnostik,* and insists that an individual's protocol should be interpreted within the framework of his *Erlebnistyp,* that is, his introversive, extratensive, or ambiversive orientation, his "experience-balance" in responding to stimuli from within and/or stimuli from without. Rorschach found that *Erlebnistyp* was reflected in the ratio between movement and color responses (M : sumC).

The scoring symbol M denotes, according to Rorschach, "Form perceptions plus kinaesthetic factors"; (29, p. 25) and he adds that "The more kinaesthesis, the more stable the affectivity" (29, p. 26). "Color responses," he continues, "have proved to be the representative of the affectivity and the rule is, the more color in the test, the greater the emotional instability of the subject" (29, p. 76).

A person with more M than C in his record would be designated the M-type (introversive), described by Rorschach (29, p. 81) as having the following characteristics:

1. Predominance of personalized productivity.
2. Intensive rapport.
3. Stable affect and motility, awkwardness, insufficient adaptability to reality and insufficient extensive rapport.

A subject with more C than M would be designated the C-type, more extratensive than introversive. The general characteristics of this type, Rorschach (29, p. 83) lists as follows:

1. The urge to live in the world outside oneself.
2. Restless motility.
3. Unstable affective reactions.

Klopfer (21) has pointed out the necessity of taking into consideration other factors in the record before venturing more than a simple statement of the subject's "natural inclinations" as reflected in his M : sumC ratio. These additional considerations give information as to whether the individual follows these inclinations or is in a state of conflict over them.

For the purpose of the present study, we may confine ourselves to an analysis of the extent to which lability and stability of the M : sumC ratio exists within the framework of our experimentally varied testing situations.

Rorschach recognized the existence of temporary variations in number of M and C in a subject's record if the mood of the subject shifted between dejection and elation; however, he stated that despite the variation in absolute number, "the proportion between them (M and C) changes little or not at all" (29, p. 94).

In the 108 records produced by our thirty-six subjects, variation in absolute number ranged from 0 to 19 on the M side and from 0 to 13 on the sumC side. The proportion between the M and C was exceedingly shifty; and, in addition, there were numerous actual shifts of weight in the ratio from one side to the other. The number and direction of these variations in Experience Balance are summarized in Table I.

From Table I it can be seen that the greatest stability of M : sumC ratio exists between administrations I and II; even here, 31 per cent of the cases shifted from

one balance to another; i.e., there was variation *within* approximately a third of the persons tested, as measured by an individual's production of more movement than color responses in one situation, more color than movement responses in another situation. A close second in rank order of stability is that between records obtained by administrators B and C, wherein 67 per cent maintained a constant balance.

The greatest number of shifts in balance occurred equally between Administrations II and III and between Administrators A and B. In both these categories, exactly half the cases maintained a constant balance and half shifted in one direction or another.

In reading Table I, it must be remembered that a shift, from M to C for example, means that the subject actually showed more M than C under the first condition considered and more C than M under the second condition listed.

A first glance at the totals in the columns labeled "Per Cent Stable" and "Per Cent Unstable" might suggest that approximately half the subjects maintained a constant stability throughout the experiment regardless of sequence, administrator, or type of administration. Such a conclusion cannot be drawn from Table I, however, since stability therein is noted within, but not between, the variable situations.

An examination of the data reveals that only fourteen of the thirty-six cases maintained a constant balance on the M: sumC ratio throughout all variations in testing situations. Of these fourteen, ten held a balance weighted on the M side of the ratio, four showed consistent weighting on the sumC side, and none maintained ambi-equality of weighting. One may, therefore, conclude that 39 per cent of the subjects retained a stable Experience Type throughout the experiment, approximately three fourths of whom were introversively oriented, a quarter of whom were extratensively oriented.

These conclusions, then, would support the prediction stated earlier in this study that approximately a third of the subjects would produce stable records. However, on the basis of Rorschach's description of the M-type and C-type personalities, noted earlier in this discussion,

TABLE I

NUMBER AND DIRECTION OF SHIFTS IN BALANCE ON M: sumC RATIOS

Situation	Frequency and Direction of Change								
	M to C	C to M	A* to M	A to C	M to A	C to A	Stable	% Stable	% Unstable
I to II	1	2	0	3	4	1	25	69	31
I to III	5	4	2	3	3	0	19	53	47
II to III	6	6	1	4	1	0	18	50	50
A to B	7	2	1	2	4	2	18	50	50
A to C	8	2	3	2	1	1	19	53	47
B to C	2	4	2	4	0	0	24	67	33
o to +	2	5	2	1	2	4	20	56	44
+ to −	4	5	3	2	2	0	20	56	44
o to −	2	7	1	2	1	2	21	58	42
Totals	37	37	15	23	18	10	184	57	43

* The symbol A is used to denote ambi-equality of M and sumC.

one would have suggested that the fourteen subjects with stable records would *all* have been of the M-type. Table I shows us that A-type, M-type, and C-type personalities in one situation can become another personality type in another situation. The additional analysis of the data shows us that both M-type and C-type personalities are capable of maintaining their personality type in the face of varying situations, but that the odds are three to one that the subject who maintains such constancy will be an M-type personality.

D. Statistical Treatment of the Data

In order adequately to test the principal hypothesis of this experiment concerning the influence of negative and positive rapport conditions on Rorschach performance, and the numerous collateral hypotheses—the influence of successive administrations and different administrators, and the kind and amount of variation in the numerous scoring categories—one must look to statistical analysis.

To justify the greatest confidence in probabilities obtained from analysis of the data, *t*-ratios were computed for the numerous pairs of series offered by the data.[1] The particular *t*-ratio used was Fisher's formula as described by Lindquist (26, p. 58 ff.) under the title "The Significance of a Difference in the Means of Related Measures":

$$t = \frac{M_o - M_H}{\frac{\text{Sum } d^2}{n(n-1)}}$$

These *t*-ratios (a total of 243) and the interpretations of them are presented in the following pages.

In working the *t*-tests, data were fractionated in the following manner: All tests given during the first session were grouped, regardless of who gave them or of what kind of affective tone marked the administration; this procedure was followed for the second and third administrations as well. All tests given by Administrator A were grouped, regardless of whether they were administered first, second, or third in the series, regardless of whether the tone of administration was positive, negative, or neutral. A similar grouping was made for tests given by Examiners B and C. All tests given during a positively loaded test-situation were grouped without respect to who administered them or to the serial number of the situation. Identical groupings were made for tests administered in the negative and the neutral situations.

This combination of heterogeneous data, of course, introduces a special problem in the interpretation of the results of the *t*-tests. If variance is found due to administrations, this variance is operating when testing the differences between administrators and between types of administrations. Where significant variance exists only within one of the three major groups of variables, this problem loses some of its importance. If, however, variance is shown to exist as a function of the repetition of the test, this variance will operate to enlarge the sigma, hence enlarge the denominator, of the *t*-tests between other variables. The influence would produce spuriously

[1] Cognizance is taken here of the criticism to which this study is subject as a result of the practice of treating scoring categories as unique parts of a whole without reference to the whole of which they are an integral part. The writer can think of no other method of experimental approach to the investigation of the Rorschach and the personality factors it is judged to measure except by analysis of the separate elements which contribute to the whole. Recognizing the imperfection of the approach, the plea is offered that it remains the best of all possible existing approaches.

low *t*-ratios among the other groups of variables.

This weakness of the *t*-test, which inevitably exists when data are fractionated and combined as in the present study, is not a serious drawback if one bears in mind that *t*-ratios approximating zero (say, under 1.000) throughout one whole set of data strongly indicate no real differences; if one remembers, also, that any significant or very significant *t*-ratio in one of the three groups of variables within a set will contribute to spuriously low *t*-ratios among the other groups, and that, therefore, a marginal *t*-ratio would be significant if the variance due to other factors were partialed out.

One other element of the statistical treatment demands discussion: Computation of *t*-ratios has been made on mean frequencies of scoring categories regardless of whether these scores are expressed as percentages of the total responses or as absolute numbers. While the percentages are not seriously affected by the variation in total number of responses, the absolute numbers are. For example, if the positive administration produces more responses than the negative, there will not necessarily be any alteration in the per cent of whole responses; there will, however, necessarily be an increase on the positive administration in the frequency of certain of the response determinants. On a statistical level, one is eager to discover *which* determinants increase; on an interpretative level, one attempts to explain *why* one determinant increases and another does not, whether the increase is a concomitant of increased responses or not.

If there were a significant difference in number of responses from the first to the second to the third administration, any differences in tabulated scoring symbols would have to be held suspect as an artifact or the frequencies would have to be reduced to ratios. However, no such differences in R were found to exist.

E. Statistical Findings

Number of responses. The greatest constancy in mean number of responses proved to be that between administrations. It is interesting to note here that Rorschach literature is consistent in reporting that the normal number of responses is around thirty with normal ranges reported from fifteen to seventy-five.

Hermann Rorschach: Normal subjects generally give from 15 to 30 responses, rarely less than 15, often more than 30 (29, p. 21).

Klopfer and Kelley: The range of responses found most frequently in all large-scale investigations of adults seems to be between twenty and forty (21, pp. 208-209).

Brussel and Hitch: Normal is around 34 . . . . Normal range is from 25 to 75 (6, p. 3).

Bochner and Halpern: The average (number of responses) seems to fall between 20 and 50 (5, p. 71).

Charlotte Buhler gives a minus weighting of 3 for fewer than twenty-five responses (7, p. 11).

As an aside, one wonders if these published means and ranges are not a bit high. Only on the positive administration and under C's administration do our subjects reach or exceed means of thirty responses.

Further in this vein, it should be noted that the previously cited study conducted by the Army Air Forces (32) shows the average number of responses for servicemen undergoing classification testing to be only twenty. Perhaps this particular Army test situation stimulated an inhibition of responses. Perhaps, too, the special population of the present study experienced a special kind of inhibition of response.

Perhaps, on the other hand, the figure thirty is not the true mean for the population as a whole, is not the appropriate cut-off point for a normal number of

responses. Possibly separate norm-groups should be established for various segments of the population.

Although there is an increase in the mean number of responses on both the negatively and positively loaded administrations over the neutral administration, with the positive eliciting more than the negative, the variability within each type of adminisration is large enough to keep the differences below the critical cut-off point for statistical significance higher than the 10 per cent level of confidence which exists between the positive and the neutral administrations.

Thre is a very significant difference, however, between the number of responses elicited by Administrators A and C. Whether giving a first, second or third administration, whether being negative, positive, or neutral, Administrator C obtained, on the average, a sufficiently larger number of responses virtually to rule out chance as a cause of the difference (1 per cent level of confidence).

Average time per response. Very significant differences were found in the average response time between the first test and both the second and the third tests. These differences can be explained rather easily in terms of familiarity with the task. On being first confronted with the Rorschach plates, the subjects uniformly spent more time formulating their concepts than they spent on subsequent reassociation with the stimuli. The reduction in response-time is uniform in direction from beginning to end of the series, despite the lack of a significant difference between the second and third administrations.

There is a consistent tendency toward a difference in response times between Administrator C and the other two administrators; in fact, the difference between A and C is significant at the 5 per cent level when rounded to two decimal places. Apparently the subjects, on the average, lingered longer over their responses when giving them to Administrator C.

Crude control. F% is considered to be one of the key determinants in a Rorschach protocol. It is the scoring symbol used to indicate the percentage of responses determined solely by the shape, contour, or outline of the blot. Briefly, it is interpreted as a measure of the control exercised by the subject. Buhler and Lefever (8) designate it as a measure of rational functioning. Excessive F% (50 per cent or above) is considered to be an indication of undue constriction, or an accompaniment of depression or anxiety.

While the means of the various measures of our subjects stay well under the 50 per cent upper limit for normalcy, we may speculate that significant increases in F% mean increased depression, anxiety, or constriction. No such significant increases are found as a function of the number or the type of the administration; however, it is interesting to observe that the lowest mean is found for the positive administration and the highest for the negative.

Again we find significant differences among administrators, A differing at the 1 per cent level with C and at the 5 per cent level with B. There is no difference between B and C.

Refined control and its components. The FK + F + Fc per cent is another measure of control; however, it is distinguished from F% alone by the quality of the control it reflects. Klopfer (21) has called the latter "crude control" and the former "refined control." Separately considered, the determinant FK (vista concept) is usually interpreted as reflecting introspection or insight; Fc (texture concept), as reflecting tact or sensitivity.

Although there is not a critically significant difference in the amount of refined control displayed in the various types of administration, the means show a consistent increase from the lowest mean on the positive administration to the highest mean on the negative administration, paralleling the tendency on F% or crude control.

A significant difference between the first and the third administration suggests several alternative possibilities. The mere repetition of the Rorschach probably brought about a degree of loss of spontaneity in responding to the stimuli. This was replaced successively more frequently by increased refined control of reactions to the stimuli.

While there is practically no difference between Administrators A and B in the average percentage of responses with this combination of determinants, Administrator C differs at the 10 per cent confidence level from A and differs significantly from B.

Further light may be thrown on the factors influencing these differences in the combination-score by inspecting the components that contribute to it. F% has already been analyzed. The raw data show a low frequency of FK and Fc responses. The responses, however, occur with adequate frequency and with sufficient approach to normality of distribution to permit application of the *t*-ratio statistic.

A *t*-ratio for Fc between Administrators B and C was not computed since the difference between means of .05 is so small as to insure a *t*-ratio of approximately zero. Despite the low frequency and small differences in means between A and the other two administrators, the variability was sufficiently small to produce a significant difference between A and C, and a very significant difference between A and B. The latter two administrators consistently elicited more Fc responses than did A. This observation would tend to rule out Fc as a contributing factor to the significant difference between B and C on the FK + F + Fc percentage. Considering, too, the insignificant *t*-ratio between B and C on F%, one might suppose that despite very low frequency and apparent scatter, the FK responses are the greatest single contributing factor to the FK + F + Fc per cent difference between B and C; however, this convenient supposition is not supported by the data, as will be subsequently shown.

It is interesting to note that F% (crude control) was sufficiently higher for A than for B or C to produce significant and very significant differences respectively; whereas, Fc (tact or sensitivity) was sufficiently *lower* for A than for B and C to produce very significant and significant differences respectively. It is not surprising, then, that when these two determinants, F and Fc, are summed with a third determinant FK the differences between A and both B and C disappear but the difference between B and C is exposed.

The scoring category FK is used for concepts incorporating vista, distance, perspective. There are no significant differences in FK on the successive administrations, although the means reflect a consistent decrease in the use of this determinant. Nevertheless, no inferences can be made from such data.

The lack of significant differences among administrators also permits no interpretation, inference, or conclusion. We may only say that our analysis of FK offered no clarification of the questions growing out of the findings for FK + F + Fc per cent. This latter category, then,

appears to be a unique whole, not paralleling any one of its three components in its stability or variation under the varied situations.

The affective loading of the administrations had a significant degree of influence on the scoring category FK. Both the negative and the positive administrations show more mean FK than does the control administration. The difference between the neutral and the positive administrations is significant at the 5 per cent level of confidence.

The symbols K and k. Responses classified as K are interpreted as a sign of "free-floating" anxiety or of insecurity. Douglas (10) found the use of this type of response serving "as conversational material until something further can be found."

K responses remain quite stable through the successive administrations. There is a noticeable difference, however, in the amount of K elicited by the different administrators, B and C differing at the 10 per cent level of confidence.

Like K, responses classified as k are interpreted as reflecting feelings of anxiety, insecurity, or inadequacy; however, the two categories are differentiated by the control factors implicit in a k response, absent in a K response.

The low frequency of the determinant k is reflected in low means for the various measures. Neither the variation of administrator nor the type of administration affected this scoring symbol to a degree worth noticing. There is a consistent, though insignificant, reduction in the appearance of this determinant on the successive tests, each administration eliciting somewhat fewer such responses than the preceding.

The C′ determinant. Our thirty-six subjects produced, on the average, about one and a half achromatic color responses (C′) per record without variation from one administration to the next. There was variation in the C′ responses with type of administration. The increase from an average of one C′ on the control Rorschachs to an average of two on the positively loaded Rorschachs boosts the confidence level between these variables to 10 per cent. Administrator A elicited the lowest frequency of C′ responses, differing from both B and C at the 5 per cent confidence level.

Movement responses. There are no statistically significant differences higher than the 10 per cent level of confidence in the number of Fm, inanimate movement, responses produced under any of the three variations of the testing situation. The constancy in number from one administration to another is so great as to preclude the necessity of computing *t*-ratios.

Mean differences, however, exist as a function of the variation in type of administration, the *t*-ratio between the control and the positive administrations exceeding the 10 per cent level of confidence that a real, not chance, difference exists.

The animal movement responses, FM, vary little or not at all on the successive administrations, and they fail to vary significantly with variation in type of administration. However, there are differences at the 5 per cent level of confidence between Administrators A and C, and at the 1 per cent level between B and C, Administrator C eliciting the larger number of responses in both comparisons.

The frequency of responses classified as M, human movement, varied sufficiently to reach the 5 per cent level of confidence between one pair of variables and the 10 per cent level between four other pairs.

Human movement is consistently interpreted as a sign of inner adjustment or equilibrium, a capacity for the absorption of emotional stimuli whether originating from within or without.

An excess of M on the positive administration over the neutral created a difference significant at the 5 per cent level of confidence, and the positive administration shows a difference from the negative at the 10 per cent level.

Administrator C provoked the highest mean frequency of M responses, differing from both A and B at the 10 per cent level of confidence. Further, the mean number of M on the third administration was sufficiently below the number on the second to give a *t*-ratio significant at the 10 per cent level.

Emotionality. The C-type personality was discussed earlier in this paper, and the relationship of sumC to M was defined. Not forgetting the importance of this inter-relationship, let us look at the meaning of color responses per se in Rorschach records. Color is considered to be the determinant of emotionality. Lack of color is considered an indication of constriction; excessive use of color is interpreted as emotional lability or overreaction to external emotional stimuli.

The use of color varied little on successive administrations. There was an infinitesimal difference between the subjects' sum of color responses on the neutral and the negative administrations; in fact, no significant differences occurred among any of the types of administration.

The really important differences in the use of color responses occurred with the variation of examiners. Administrator A elicited significantly fewer color responses than C and very significantly fewer than B.

Content categories. Normally, the sum of animal and animal-detail concepts constitutes approximately 50 per cent of the content of Rorschach concepts. Our subjects linger close to this norm. An increase in A% is usually interpreted as evidence of stereotypy of thinking or a repression of intellectual activity; a decrease in A%, contrarily, is widely interpreted as less confinement to the obvious, a broader range of thinking and/or interests.

There are completely unimportant differences in A% on successive administrations of the test. There is no appreciable difference in A% from one examiner to another; although, again, the differences between Administrator A and the other two examiners show a stronger tendency toward significance than does the difference between B and C, that between A and C closely approaching the 10 per cent level of confidence.

The percentage of A% responses on the affectively loaded administrations differs significantly, with the negative administration producing the percentage higher than that on the positive administration. Also, the neutrally conducted tests deviate from the positive with a difference at the 10 per cent level of confidence.

There is agreement among users of the Rorschach that the number of animal responses should and do exceed the number of animal-detail responses in a normal, healthy record. Therefore, it was considered advisable to break down the A% scoring symbol into its component parts in an effort to examine the circumstances under which both animal responses and animal-detail responses may vary or remain stable.

Examination reveals no significant differences in the number of whole animal responses elicited in the various situations. Even the means do not show a dif-

ference worth comment. Certainly this analysis of the number of whole animal responses does not help appreciably to clarify the differences obtained in the A%. A look at the means and critical ratios for animal details is indicated; however, one must not overlook the fact that absolute number frequencies introduce different considerations from those under consideration when one is speaking of percentages.

In all measures, the number of A responses exceeds the number of Ad responses, indicating that our subjects on the average maintain the healthier balance regardless of variation in testing situations. With both symbols there are infinitesimal differences in frequencies from one testing situation to the next. Among administrators, C elicits the highest frequency for both A and Ad responses. Among the types of administrations, the control Rorschachs show the lowest mean for both scoring categories.

Differences at the 5 per cent level of confidence exist between Administrator A and both B and C, with A gleaning significantly fewer Ad responses than B or C.

As with other scoring categories, the number of A and Ad responses by themselves are considered to have little meaning. Whereas, an increase in Ad may indicate a tendency toward a critical attitude, this interpretation of excessive Ad in a record is made only if the sum of animal-detail and human-detail responses exceeds the sum of whole animal and whole human responses. This observation calls for an analysis of the number of human and human-detail responses.

Several measurable differences in number of whole human responses exist. The frequencies are exceedingly stable from the first through the third administration. While Administrator C provoked a larger mean number of human responses than did either A or B, only the difference with A is great enough to produce a *t*-ratio significant at the 10 per cent level of confidence.

The means of the types of administration show an increase in human concepts on both the negative and positive administrations over the number on the control administration, with the difference between the neutral and positive rising to the 10 per cent level of confidence.

Again, the frequency of whole versus part concepts of human figures is a more important consideration than the mere number of either.

A comparison of means reveals that the records of our subjects, regardless of variables, maintain a healthy preponderance of H over Hd responses.

Administrator C evoked a greater number of Hd responses than either B or A, and the mean of C's administrations differs significantly from that of A.

Popular responses. The means of the popular responses elicited by the various examiners are identical. There is a slight but steady increase in mean P from the first to the second to the third administration, which may easily be explained in terms of increased familiarity with the material through mere repetition.

The difference between the means of the negative and positive administrations reaches the 10 per cent level of confidence, the permissive situation stimulating the higher number.

Use of ground as figure. Neither repetition of administration nor affective loading of administration produced any difference in the number of white-space, S, responses on the Rorschach records. Neither are Administrators A and B essentially different in the amount of S they provoked on the protocols. Administrator C, however, differed from both A

and B in eliciting a higher frequency of white-space responses, and differed from A at the 5 per cent level of confidence.

An earlier discussion of additional responses explained why they were not incorporated into the analysis of results on this study. An exception should be made, however, of the white-space scoring category. Among normal or near-normal subjects, white-space responses occur most frequently concomitant with reaction to inked areas of the plates. Arbitrary scoring methodology requires that, in such mixed concepts, the inked area be scored as the main response location and the white space be scored as an additional response location. Consequently, many clear-cut reactions to white space are buried among the additional responses. Better to understand the actual amount of and variability in the use of white space, it was decided to sum the absolute number of main and additional S responses and analyze the distribution of this total reaction to ground.

The total use of white spaces by our subjects varies significantly with type of administration and among administrators. The smallest amount of total S + (s) was elicited on the neutral or control administrations, the most on the positive administrations. While the actual mean difference between neutral and negative administrations is small, the direction of change is so consistent, the variability within the columns so small, that the difference reaches the 5 per cent level of confidence. The higher mean difference between neutral and positive administrations is accompanied by sufficiently increased variability to hold the difference down to a 10 per cent level of confidence.

Among administrators, there exist real differences in the amount of sum-S produced, regardless of number or type of administration. A evoked the least, and C the most, use of white space. The difference is significant at the 5 per cent level of confidence. C also exceeded B in elicitation of total S responses with a difference significant at the 10 per cent level. The most significant difference, however, is between administrators A and B, with a *t*-ratio at the 1 per cent level of confidence.

Percentage of responses to all-color cards. Color, as previously stated, is interpreted as a correlate of emotionality. Since cards VIII, IX, and X on the Rorschach test are composed exclusively of bright-color inks, they are construed to have particular significance in reflecting emotional components of a subject's personality. The normal expectancy of responses on these all-color cards is approximately 40 per cent of the total responses.

Our subjects consistently approximate normal expectancy, with little variation from one administrative situation to another. The largest difference occurs between the neutral and negative administrations, where the *t*-ratio rises to a 10 per cent level of confidence.

Apparently the over-rejecting situations brought about a measurable decrease in response to external emotional stimulation, creating a dampening effect on the subjects.

Manner of approach. The number of whole responses in a Rorschach record provides two separate clues for interpretative analysis: (1) a clue as to the percentage of responses which are based on incorporation of the total blot-stimulus in the concept-formation; (2) a clue as to the ratio between whole responses and movement responses, i.e., the W : M ratio.

The first of these clues will be explored later upon the presentation of means and *t*-ratios for W%. The second requires the reintroduction of the means for the scoring symbol M.

There is a steady decrease in the mean number of whole responses with each repetition of the test, the difference between number of W responses on administrations I and III being significant at the 5 per cent level of confidence.

If the W : M ratio much exceeds 2 : 1, the interpretation is usually made that the subject is over-extending himself; i.e., he is striving beyond his means of achievement. It is not considered a healthy sign. Let us compare our W : M mean ratios for any evidence of this imbalance.

One ratio that was obtained on the second administration achieves the ideal relationship of 2 : 1. With one exception, all other ratios show heavier W weighting; however, the ratio of 3 : 1 is not sufficiently in excess of the ideal to merit interpretation of a significant imbalance. The ratio on the neutral administration of 4 : 1, however, is twice the expectancy ratio. On the basis of Rorschach literature, one would have predicted a 2 : 1 ratio on this control administration, regardless of what other predictions attended the other administrations.

The really crucial interpretative factor related to whole responses is the scoring symbol W%—the percentage of total responses which use all of the inked surface of the cards within the framework of a single concept. "In a normal record," say Brussel and Hitch, "about 25 per cent of whole answers are expected" (6). The Individual Record Blank for scoring Rorschach protocols, developed by Klopfer and Davidson, and in wide use, shows a normal range of W% extending from 20 to 40. Every mean in the present study exceeds not only the figure given as a normal mean but also the figure given as the upper limit in the range of normal expectancy for W%. One might assume that the stated norms are in error; however, it seems less arrogant to consider the possibility that the subjects in the present study deviate from the population as a whole somewhat.

Examination supports the earlier finding of reduction of whole responses as a mere function of repetition of the test. When working with percentages rather than absolute numbers, this reduction in tendency is more dramatically revealed. While the decrease in W% is insignificant from the first to the second administrations, the decrease is significant at the 1 per cent level between the third and both the first and the second administrations. C's protocols show the lowest mean percentage of whole responses and differ at the 5 per cent level of confidence with A's records, which show the highest mean W% among administrators.

While none of the differences in means of varied types of administration are statistically significant, it is interesting to observe that the negative administration shows the highest percentage of whole responses and the positive situation the lowest percentage.

As W% decreases, it is mathematically certain that other location sub-divisions will increase. The nature of these increases can be explored only by analyzing the frequencies in the data of large details, small details, and tiny or unusual details as they are employed by the subjects in concept-formations.

One clear-cut difference exists between Administrators A and C. Administrator A elicited the highest percentage of W% responses and the lowest percentage of D% responses; in both scoring categories, A and C differ at the 5 per cent level of confidence.

No other variation in testing situation creates differences of statistical significance. Nevertheless, consistent tendencies are revealed by the means of the various

types of administrations. The negative administration provoked the lowest mean D% and the positive administration the highest. The *t*-ratio between them barely misses reaching the 10 per cent level of confidence.

The location scoring symbol D is used when subjects form a concept using a large, insular part of the blot, frequently perceived as a separate entity by normal subjects. Between 45 per cent and 55 per cent of total responses are expected to be D responses.

Not until the third administration did the subjects show any measurable increase in the use of small usual details in the blots. The production of concepts based on these small areas remained constant from the first to the second test but differed between the second and the third to a degree significant at the 10 per cent level of confidence.

The average intelligent subject is expected to have from 5 per cent to 15 per cent of his responses in the d% category. The third administration, then, most nearly provoked d% responses up to a minimum of expectancy for our intelligent subjects. Decreased d% is not generally interpreted as having any particular significance, and increased d% is ignored unless it exceeds 15 per cent of the total responses.

Although there is a lack of significant differences between the various administrators, Administrator C has the highest mean d% within the administrator group, just missing 10 per cent *t*-values with both A and B.

Dd + S% is a sort of waste-basket category for all responses based on blot areas not defined as W, D, or d. Such responses may include tiny details, inside or edge details, unusual combinations of large or small blot areas, and the use of ground as figure. White-space responses were analyzed earlier by absolute number because of their special qualitative meaning in Rorschach interpretation. Their tabulation in combination with unusual details is a necessary part of the process of recording percentages of responses by location areas alone.

There is complete agreement among Rorschach authorities that absence of Dd + S% is normal. There is, however, some difference of opinion as to how high this percentage may go before the upper limit of normalcy is exceeded. The range of estimates is from 2 per cent to 10 per cent, with the consensus approaching the larger figure. In any event, the records of normal persons rarely exceed 10 per cent in this category. On all variations of the testing situation, the subjects in this study remain under this 10 per cent maximum; however, the ceiling is approached in the means of third administration and of the positively administered tests. None of the differences between means are statistically significant.

While Dd + S per cent in excess of 10 per cent is interpreted as an abnormal sign, even an excess of 5 per cent in the category is, by some authorities, considered to be evidence of a pedantic approach to the test (and, by extension, to the environment and to external stimuli, in general). All nine of our means exceed 5 per cent. The foregoing interpretation is more readily made if the higher percentage of Dd + S responses occurs in conjunction with high W%, a fact which we have already observed to exist with our group.

Summary of t-ratios. This completes the statistical analysis of the data on the numerous Rorschach scoring categories. The somewhat lengthy analysis of means and *t*-ratios by Rorschach scoring categories seemed the most clear-cut way of presenting the data of this study. How-

ever, that procedure inevitably produces some confusion as to how many statistically significant *t*-ratios were found and where, how frequently Administrator A or administration + had the highest means and on which types of response. These data are presented in summary form in Tables II to IV.

which each variable, within each of the administrative groups, achieved the highest, middle, and lowest means. It will be noted that the positive administration and Administrator C reveal almost identical frequency patterns. No other pairs show this marked parallel tendency.

As Table IV reveals, there were, at the

TABLE II

RANK ORDER OF MEANS, ACCORDING TO TYPE OF RESPONSE

Type of Response	Administration			Administrator			Type of Administration		
	I	II	III	A	B	C	+	−	o
R	2	1	3	3	2	1	1	2	3
T	1	2	3	3	2	1	1	3	2
F%	3	2	1	1	2	3	3	1	2
FK	1	2	3	2	3	1	1	2	3
K	1	2	3	2	3	1	1	2	3
Fk	1	2	3	2	3	1	1	3	2
≠F	3	1	2	1	3	2	2	1	3
FC	1	3	2	3	2	1	1	2	3
CF	1	2	3	3	1.5	1.5	1	3	2
C	1	3	2	1	2	3	1	3	2
FK+F+Fc/R	3	2	1	2	1	3	3	1	2
M	2	1	3	2	3	1	1	2	3
sumC	1	3	2	3	1	2	1	3	2
FM	2	1	3	2	3	1	1	3	2
Fm	2.5	2.5	1	3	2	1	1	2	3
Fc	3	2	1	3	2	1	1	3	2
C′	1	2	3	3	2	1	1	2	3
A%	2	3	1	1	2	3	3	1	2
P	3	2	1	2.5	2.5	1	1	3	2
H	2	1	3	2	3	1	1	2	3
A	2	1	3	2	3	1	1	2	3
Hd	3	1	2	3	2	1	1	2	3
Ad	3	2	1	3	2	1	2	1	3
CR%	2	3	1	3	1	2	2	3	1
W%	1	2	3	1	2	3	3	1	2
≠W	1	2	3	3	2	1	1	2	3
D%	3	2	1	3	2	1	1	3	2
d%	2	3	1	3	2	1	2	3	1
Dd+S%	3	2	1	3	2	1	1	2	3
S	1	3	2	3	2	1	1	2	3
S+(S)	1	3	2	3	2	1	1	2	3

Table II serves to show the administrative circumstances under which the means of each of the thirty-one Rorschach response categories, herein analyzed, held the highest, middle, or lowest numerical position within its group.

Table III shows the frequency with 1 per cent level of confidence, a total of ten differences, over four times as many significant *t*-ratios as one would expect by random sampling; four of these were a function of repetition of the test, six a function of examiner differences. Administrator A differed very significantly from

B on three of the scoring categories and from C on two of the measures. B and C differed only once at this level of significance.

At the 5 per cent level of confidence, one would expect twelve significant differences by random samplings; whereas, we obtained twenty-one additional differences at this level alone. Combined with the ten *t*-values at the 1 per cent level, our total is thirty-one *t*-ratios equal to or above the value required for confidence at the 5 per cent level. This figure is almost three times chance expectancy. Successive administrations accounted for two of the significant differences, and type of administration for four. The fifteen differences among examiners point up the greater frequency between A and C than between either A and B or C and B. Three of the administrator differences were between A and B, one between B and C, but eleven existed between A and C. Combining differences at the 1 per cent and 5 per cent levels, A differs six times with B and thirteen times with C, while B and C differ a total of twice.

TABLE III

FREQUENCY OF MEAN POSITION, ACCORDING TO ADMINISTRATIVE VARIABLE

Administration	Highest Mean	Middle Mean	Lowest Mean
I	13	8½	9½
II	7	15½	8½
III	11	7	13
A	5	8½	17½
B	3½	19	8½
C	22½	3½	5
+	23	4	4
−	6	14	11
o	2	13	16

Differences (5 per cent level) resulting from variation in the emotional tone of the testing situation occurred twice between the neutral and negative administrations, once between the positive and neutral, and once between the positive and negative administrations.

Clearly, the largest number of significant and very significant differences at the 1 per cent and 5 per cent levels were the result of examiner differences; and, within our group of examiners, A was the greatest contributing factor to frequency of differences. The negative and neutral administrations appear with equal frequency (three times each) as contributing factors to significant differences in type of administrations. The positive administration is a factor in two of the four differences.

Table IV summarizes the seventeen *t*-ratios at the 10 per cent level of confidence. This figure brings to forty-eight the total number of *t*-values reaching or exceeding the 10 per cent confidence level, a total exactly twice as large as the twenty-four one would expect by random sampling.

At this level of significance, 10 per cent, the variations with type of administration predominated, with nine differences tending toward significance. Two differences were a function of repetition of the test, and six occurred with change of administrators. Administrator C was a party

TABLE IV

NUMBER OF MEANINGFUL *t*-RATIOS BY VARIABLES

Confidence Level	I II	II III	I III	A B	A C	B C	+ −	+ o	− o	Cumulative Total
1%	1	0	3	3	2	1	0	0	0	10
5%	0	1	1	3	11	1	1	1	2	31
10%	0	2	0	0	3	3	2	6	1	48

to all of these variations, differing three times with A and three times with B. Of the nine differences occurring with variation in the tone of the administration, the positively loaded situation differed six times with the neutral administration and twice with the negatively loaded situation. The negative and positive administrations showed differences at this level of confidence only twice, the negative and neutral only once. Summarizing, the positive administration was a factor eight times, the neutral seven times, and the negative three times.

Summary of results by scoring categories. A few of the Rorschach scoring categories proved exceedingly stable, showing no variation from test to retest, from examiner to examiner, from positive to negative to neutral administration. These stable categories are the following: Dd + S per cent, Fk, and number of A responses.

The majority of the Rorschach scoring categories showed measurable variability as the testing situation varied. The degree of variation, together with the administrative variable which produced the variation in each shifting category, was as follows:

a. R: There was one difference at the 1 per cent level of confidence between administrators and one difference at the 10 per cent level with variation in types of affective loading of the administrative situation.

b. T: There were two differences at the 1 per cent level in average time per response on the successive administrations and one difference at the 5 per cent level between administrations.

c. M: This scoring symbol proved to be quite variable, with one difference at the 5 per cent level and another at the 10 per cent level as a function of the varied affective loading of the testing situation; two differences occurred at the 10 per cent level with variation in administrators, and one difference at the 10 per cent level grew out of mere repetition of the test.

d. FM: Neither successive administrations nor variation in affect in the testing situation influenced FM. Change of administrators, however, produced one difference at the 1 per cent level, another at the 5 per cent level.

e. Fm: This scoring symbol was resistant to change with repetition of the test or with change of examiners. It varied only with the variation in type of administration, and then only at the 10 per cent level of confidence.

f. K: Administrator differences produced one difference in K at the 10 per cent level of confidence. No other variables in the testing situations measurably influenced this determinant.

g. FK: This response determinant also remained stable through successive tests and despite variation in examiners, but it showed a 5 per cent level of significant difference with variation in affective loading of the testing situation.

h. Fc: Examiner difference again is evident with this determinant, there being both a 1 per cent and a 5 per cent difference among administrators. Neither retesting nor affective loading affected the mean number of Fc produced.

i. C′: This scoring symbol varied once at the 5 per cent level with repetition of the test, twice at the 5 per cent level with change in administrators, and once at the 10 per cent level with variation in tone of administration.

j. sumC: The composite score sumC remained stable from one test to the next and despite variation in affective loading of the testing situation; however, the mean scores varied significantly twice with change of administrators, once at the 1 per cent level, once at the 5 per cent level.

k. F%: This factor also resisted change with number and type of administration; however, it, too, varied with the examiner once at the 1 per cent level, again at the 5 per cent level.

l. A%: The difference in the type of administration produced two differences in A%, one at the 5 per cent level, the other at the 10 per cent level. Neither administrators nor successive testing influenced the means significantly.

m. P: The number of popular responses remained somewhat constant throughout the experiment. One difference, at the 10 per cent level, occurred with variation in the affective loading of the testing situation.

n. FK + F + Fc / R: The affective loading of the administration did not influence this composite scoring factor; however, both successive testing and change in administrators brought about one difference each at the 5 per cent level of confidence. Examiner differences accounted for an additional difference at the 10 per cent level.

o. H: The number of human-figure responses varied at the 10 per cent level once with change in administrators and once with variation in the type of administration.

p. Hd: Human-detail responses were unaffected by retesting or by affective loading of

the test situation; however, there was one difference at the 5 per cent level with change in administrators.

q. Ad: Animal-detail responses were also impervious to the emotional tone of the administration and to successive testing, but they twice showed a difference at the 5 per cent level with variation in administrators.

r. CR%: The percentage of responses to the all-color cards remained somewhat stable, showing only a single difference at the 10 per cent level of confidence upon variation of the affective loading of the administration.

s. #W: The number of whole responses varied once at the 5 per cent level with successive testing but remained impervious to change with varied administrator or type of administration.

t. W%: W% remained unchanged with variation in emotional loading of the test situation, but showed two differences at the 1 per cent level with retesting and one difference at the 5 per cent level with variation in administrators.

u. D%: This location scoring symbol varied once at the 5 per cent level with change in examiners but otherwise remained quite unchanged.

v. d%: The percentage of small usual details remained stable except for one difference at the 10 per cent level with successive testing.

w. S: The number of white-space responses did not differ from one examination to the next but varied between examiners once at the 5 per cent level and differed once at the 10 per cent level with alteration of affective loading of the administration.

x. S + (s): The sum of main and additional white-space responses resisted change with successive administrations; however, the variation in affective loading produced one difference at the 5 per cent level. Change of administrators proved to be the most significant variable with one difference at the 1 per cent level, one at the 5 per cent level, and a third at the 10 per cent level of confidence.

F. Review of Results Related to Original Hypotheses

Let us briefly review the specific hypotheses and predictions made at the outset of this experiment and check our results against them.

1. The crucial hypothesis related to the stability of responses to the ink blots regardless of variation in environmental stimuli other than the ink blots; i.e., the person presenting the blots, the implicit attitude of the administrator, or the mere repetition in presentation of the blots. Of the twenty-three Rorschach elements subjected to statistical analysis, only three were relatively unaffected by the extra-test stimuli, by "situational" stimuli. Nineteen of the twenty-three elements entered into differences at least once at the 1 per cent, 5 per cent, or 10 per cent levels of confidence. These nineteen varying factors showed a total of forty-six differences attendant upon some variation in stimuli other than the test stimuli—the ink blots themselves. The hypothesis of stability does not hold.

2. It was predicted at the outset that the number of responses would vary with administrators, regardless of number or affective tone of the administration. The means of responses for the three administrators were, roughly, 24, 27, and 33, the difference between highest and lowest means being significant at the 1 per cent level of confidence. These findings support the prediction.

3. No consistent variation in the number of responses was anticipated from the first to the second to the third administration; and, in fact, there was none. Initial records of over seventy responses were expected to decrease with successive administrations, which they did. However, the prediction that initial records under thirty responses would show successive increase in number of responses was not supported by the data.

4. Stability as a function of individual differences reflected in M : sumC ratio proved, as predicted, to isolate approximately 30 per cent of the cases. Also fulfilled was the prediction that the more stable personality records, those of the M-type individuals, would show more situational stability than the records of the C-type persons. Thirty-nine per cent of the subjects retained a stable Experi-

ence Type throughout the experiment, approximately three-fourths of whom were M-type personalities, one-fourth C-type.

5. The hypothesis that there would be some variation in the records obtained by the three administrators, regardless of number or type of administration, was more than supported by the data. In fact, examiner differences proved the most pervasive influence in the experiment—perhaps the most dramatic finding of the study.

6. The predicted increase in frequency of k, K, c, and C′ with the negative administrations was in no particular realized. In fact, both Fk and Fc had the lowest mean frequency in the negative testing situation; the other symbols—K, FK, and C′—showed means on the negative administrations which were exceeded by the mean frequencies on the positive administrations. Apparently, as judged by these determinants, our thirty-six subjects found the positively loaded testing situation more challenging or threatening than the negatively loaded. This finding negates the hypothesis under discussion.

7. The hypothesized increase in white-space responses in the negative testing situation was somewhat supported, if one judges negative means against control means. Both main S and main-plus-additional S responses showed a mean increase in the negative administrations over the control administrations. However, the negative means of both measures were exceeded by the means on the positive administrations. Apparently the negative administrations did provoke more oppositional behavior than the control administrations; however, the positive administrations elicited even more of this type of response than either the negative or the neutral administrations, a finding which, in effect, negates the hypothesis of peak negativistic behavior on the negative administration.

8. Both crude and refined control, as measured by F% and FK + F + Fc per cent respectively, were expected to increase in the negative administrations. These predictions were supported by the data.

9. The anticipated increase in emotionality on the positive administration, as measured by sumC, was realized. There was a steady rise in mean sumC from the negative to the neutral to the positive administrations.

10. The increase in movement responses (M and FM) on the positive administrations occurred as predicted. Both scoring categories achieved their highest mean with the positive administration. The lowest mean frequency on FM occurred with the negative administration; on M, with the neutral administration.

V. DISCUSSION OF RESULTS

In the foregoing report of all of the significant findings for the experimental variables of this study—administrations, administrators, and types of administrations—there was little attempt to make qualitative evaluations or dynamic interpretations of the findings. Nevertheless, one is justified in attempting—in fact, is obligated to attempt—to relate these data to the dynamics of personality.

A. Experimental Variables

An effort will be made to summarize the effects produced upon the experiencing subjects by the different variables herein studied. Too, an attempt is made to understand the differences among the three examiners in terms of the different effects they so consistently produced in the test results. While these procedures have highly speculative components, every effort is made to hold the speculations close to the actual data, to translate into dynamic terms the actual findings of this experiment.

Influence of repetition of the test. The mere repetition of the Rorschach was the least important variable in the entire study, accounting for the smallest total number of important *t*-values, only eight in all. The categories showing change with successive administrations are, for the most part, such as can be quite easily accounted for. One is not surprised to find the average time per response consistently decreasing with repetition, this change accounting for two administrative *t*-values. On being successively confronted with the same stimuli, subjects needed no time to adjust to the stimuli; they could employ mere recall as a substitute for perception, apperception, and conceptualization. Furthermore, there is the strong possibility that the subjects, on once learning the total demand or expectancy offered by the test-challenge, simply functioned in line with our cultural speed-value, meeting the total requirements of the situation with as little expenditure of time as could be managed.

We have elsewhere indicated that our subjects frequently showed moderately compulsive tendencies. As a defense mechanism, the compulsive tendencies would be expected to be most operative when the subjects felt most challenged or threatened. One influence of repetition of the test on this dynamic pattern would be reduction of threat or challenge and consequent reduction in display of the compulsive defense. Three of our eight administration *t*-values grow out of reduced use of the whole blot in concept formation on successive tests. This phenomenon would seem to need no further comment.

Repetition of the Rorschach induced a successive and consistent increase in the use of "refined control." Again, one does not have to delve into deep dynamics of personality to understand this finding. Our intelligent subjects merely demonstrated their ability to improve their mastery over a challenging situation with each successive association with the test-task.

The remaining measurable influence of repetition on the test results is concerned with M, human movement responses. On taking the test for the second time, our subjects manifested greater equilibrium, more inner adjustment, a higher level of imaginative thinking. This is in line with the previously noted increase in security, the display of refined control over the situation which attended successive administrations. The third administration of the same test, however, failed to stimu-

late our subjects to display this sign of the superior, well-adjusted person. There is a measurable decrease in human movement responses on the third administration. One may speculate that the subjects were inhibited on the first administration, were less inhibited and more secure on the next association with the blots, hence free to demonstrate optimal capacity to absorb and creatively utilize stimuli, but were perhaps just bored with the third presentation of the same blots, bored to the point of not bothering to behave optimally.

The suggestion that boredom attended the third testing situation is somewhat supported by the spontaneous comments of many of the subjects during the third session: "These again?" "I've already done this twice." "Same old thing—a butterfly," etc., etc.

Effect of Examiner A. Among administrators, A elicited responses which reveal the following functioning of the subjects, as indicated by rank order of means: Fewest number of responses, shortest average time per response, least degree of emotional adjustment or social adjustment (FC), least amount of emotional lability (CF), smallest measure of emotionality (sumC), least expression of sensitivity (Fc), least tendency to use all possible environmental stimuli (C′), least stimulation of critical or perfectionist attitudes toward concepts (Ad and Hd), lowest frequencies in all location areas except W% wherein A's subjects show the highest mean.

The signs on which A's mean frequencies were highest in rank order suggest the following reactions to the effect of this administrator on the subjects: Greatest amount of crude control (F%), most frequent manifestation of unhealthy, uncontrolled direct emotional interaction with the environment (C), strongest tendency toward compulsive behavior (W%).

The pattern which seems to form, as one reviews these summarized effects of Administrator A on the subject's Rorschach records, somewhat resembles what one might expect of a person confronted with a threatening, frustrating situation. If this total effect may be considered a mirror of the administrator's personality, then Examiner A would be described as a cold, forbidding, frustrating, threatening figure, these personality components permeating the test-situation regardless of the deliberate and varied role-playing attempted.

There is no way of checking the validity of this thumb-nail sketch of A's personality components as reflected in her effect on Rorschach records. The writer, however, did ask two psychologically sophisticated persons, acquainted with all three examiners, to give subjective descriptions of each. Both described A as the coldest, most inflexible, and most solid of the examiners. One added the adjective "masculine" and the phrase "castrating type of female."

Effect of Examiner B. Administrator B's results fell in the middle position on rank order of means with the majority of measures analyzed in this study. The relative effect of her personality on the administrations is revealed in part, perhaps, by the very paucity of extreme positions achieved under her administrations. Following are the measurable effects: smallest amount of introspection (FK), lowest indication of maturity or creative imagination (M), and least evidence of bouyancy or free expression of primitive drives (FM).

The subjects, under B's administrations, achieved rank order of 1 on only two significant means of categories: high-

est total degree of emotionality (sumC) and most refined control (FK + F + Fc/R).

The pattern reflected in the mirror of effects suggests that B may possess personality traits that are emotionally exciting to the subjects; at the same time, this excitement apparently is well controlled and does not provoke undue anxiety or tension, neither does it stimulate intellectual activity.

It will be remembered that Administrator B is the youngest of the three examiners. She was subjectively described as the most feminine of the group, the "softest," most nearly the "ideal protecting mother-figure." One person described her as essentially "seductive."

Effect of Examiner C. Administrator C stimulated in subjects the least amount of either crude or refined control (F% and FK + F + Fc/R), the lowest measure of uncontrolled emotionality (C), and the least indication of "whole compulsion" (W%).

The means of the significant measures showing highest rank order under C's administrations are the following: Most responses with shortest average time per response, most introspection and unfixed anxiety (FK and K); best social and emotional adjustment (FC), most evidence of creative imagination (M) and of bouyancy (FM), highest indication of sensitivity (Fc), freest use of all possible stimuli (C′), most evidence of ease in relating to people (H), greatest evidence of critical or perfectionist attitudes toward productions (Ad and Hd), greatest use of all blot location areas except W% wherein C's subjects show the lowest mean.

The effect of Examiner C on the results of the Rorschach are more complex than the effects of the other examiners. One may speculate that C provided more of an intellectual than an emotional stimulation to the subjects. There would seem to be evidence of challenge with attendant anxiety, but there is also evidence of easy rapport, social adjustment, and less need for control devices in the situation with C.

C is the eldest of the three examiners. She was subjectively described as the most flexible of the three, as more feminine than A but less than B. One person described her as exuberant, "bubbling," the other as "very sympathetic."

Before leaving these speculations on the effects of the various examiners, it should be pointed out that Administrator A quite consistently approximates the pattern of the negative administrations with sixteen agreements out of a possible thirty-one. Administrator B most nearly approximates the pattern of the neutral administrations with fourteen agreements out of thirty-one. Administrator C almost duplicates the pattern of the positive administration with twenty-six agreements out of the possible thirty-one.

One may speculate that the basic personalities of the examiners, which permeated all testing situations regardless of the varied affective roles they played, are in fact related: A = negative; B = neutral; C = positive. Or one might infer that Administrators A and B, for example, in deliberately structuring the positive administrations actually created a situation somewhat like that which was consistently and unintentionally created by C in all administrations, that A and C, in attempting to be neutral succeeded only in resembling B's essential personality, and that B and C in structuring negative situations succeeded merely in resembling A's basic pattern. The former interpretation seems more likely. Whichever interpretation appears more reason-

able, the facts and figures serve further to underline the importance of examiner differences.

Effect of positive administrations. This highly permissive situation was deliberately structured to give subjects the feeling of acceptance; the atmosphere was one of approval from beginning to end. This procedure brought about significant changes in the Rorschach behavior of the subjects, as has been previously noted. Let us summarize the effect of these positive administrations:

More responses were elicited, suggesting greater productivity in the more permissive situation. These responses tended to incorporate all possible stimuli presented. There was a greater display of intellectual activity, of creative imagination, plus evidence of increased thinking of a popular or communal type. Stereotypy of thinking was markedly low during the positive administrations. There was evidence of increased resistance against feelings of personal inadequacy plus a rise in tension-tolerance and in self-questioning, or introspection. Finally, there was marked evidence of a greater ease in relating to other human beings.

Effect of negative administrations. The negative test-situations were non-permissive, rejecting; an atmosphere of disapproval prevailed throughout. The effect of this environmental tone on the Rorschach records is reflected in a definite increase in stereotypy of thinking and a decrease in thinking that is in harmony with community thinking. There is a low incidence of intellectual activity of the imaginative or creative sort. Too, there is evidence of withdrawal from emotional stimuli, as measured by a decrease in the number of responses to the all-color cards. There is a marked rise in self-questioning, or introspection, plus evidence of resistance against feelings of personal inadequacy.

Effect of neutral administrations. The chief value in the experimental design of the neutral set of Rorschachs is that of a control factor. Since these tests were administered in the standard manner, without affectively loaded pre-tests, they serve somewhat as a measure of the reactions to the Rorschach which one might expect to find under ordinary clinical situations. The results on the neutral Rorschachs, however, cannot be construed to have the same meaning as is ordinarily attributed to the results of control groups in psychological experimentation. A third of our subjects had had one affectively loaded Rorschach before experiencing the neutral situation; two thirds had had two affectively loaded experiences with the test before receiving the neutral administration. It would be folly to assume that all of these neutral administrations provided identical test-situations for our subjects, that there was no carry-over from previous experiences with the administrative tone in which the blots were presented. Nevertheless, our neutral administrations do offer a modicum of evidence as to how subjects react to the test when standard testing conditions prevail.

The neutral situation elicited fewer responses than the positive, and these responses showed higher stereotypy of content than occurred on the positive administrations. There was less interest in utilizing all environmental stimuli, diminished ease in inter-personal relationship, and diminution of intellectual creativity on the neutral—as compared with the positive—administrations. Lower tension-tolerance and less resistance to feelings of inadequacy were also apparent in the neutral, over the positive, administrations. A lower resistance reaction

was evident on the neutral administrations when compared with the negative situations. Less introspection was apparent on the neutral than on the positive administrations, and there was greater reaction to the emotional stimulation of all-color cards.

B. Implications for Further Investigation

This study does not supply the answer to a question which grows out of the findings: What kind of personality was possessed by each of the three administrators? Since administrator differences appeared as the greatest contributing factor to the variation in the Rorschach records, one would like to know what specific and measurable factors in an examiner's personality will bring about what specific and measurable factors in a subject's Rorschach protocol. One wonders if a Rorschach psychogram of the three examiners would resemble the composite mean profiles of the subjects in every way or in any way. These problems should be investigated.

The present study was designed in such a way that a subject could withdraw without suffering too great a penalty. Consequently, several of the individuals who reacted intensely to the negative administration withdrew from the project without finishing the three tests. Their records, of course, could not be included in this study. This entire experiment should be repeated with a group of subjects over whom the examiner can exercise greater control. The fact that differences were revealed as a function of the negative administration makes one wonder how much more significant these differences might have been—and what additional differences might have appeared—if the records of the intensely affected subjects could have been incorporated in the data.

The experimental design of the present study required that each subject not only experience three different affective loadings of the test situation but also experience three different examiners. Two separate investigations should be made as a supplement to the present study. It would be well to know whether the same kind and amount of fluctuation of scoring symbols and M : sumC balance would occur if a single examiner tested a group of subjects under varied affectively loaded situations or if several examiners successively tested a group of subjects under the same affectively loaded circumstances. This latter study should, in fact, be broken into at least two studies, one investigating the results of a series of positively loaded situations, and the other exploring the results of a series of negatively loaded administrations.

In order to minimize the personal equation and thereby to study in purer form the influence of negative and positive rapport factors on Rorschach performances, one could repeat this experiment with the following variations: neutral, positive, and negative administrations could be administered in group form, according to the design herein used, with directions or instructions given by a single examiner through the medium of phonographic recordings. This examiner would, of course, have to cut three records, each affectively differently loaded. While such a procedure could not effectively incorporate the particular card-sorting pre-tests used in this study as tone-setters for the administration, other equally simple devices could be created. In fact, the findings of the present study suggest that no such elaborate preamble is necessary to instill feelings of accep-

tance or rejection. A cold reception proved adequate to disturb our subjects to the point of misunderstandings and errors.

The significant findings for the group of male college sophomores justify a repetition of this entire experiment with other age groups and with female subjects. While our subjects were randomly selected, they were not selected from the population as a whole but from a rather small stratum of a very small and highly selected population—college sophomores. Our group was shown to deviate from population norms in several instances; e.g., in compulsive tendencies and pedantic approach. One may only speculate on the extent to which other findings in this experiment are a function of the select group from which the subjects were drawn. Would a similar study, with subjects representing a cross section of the population as a whole, show more or less variation, show the same or different kinds of instability?

All of these questions, and many more, must eventually be answered in the laboratory if we are to refine our use of the most popular and pragmatically valuable of the projective techniques: The Rorschach Ink Blot Test. This study is a small step in a direction which must, and certainly will, be taken by other students who realize that the practical success of of the Rorschach must be supported by laboratory verification. This area is rich in experimental problems sufficient to excite at least a generation of students.

VI. SUMMARY AND CONCLUSIONS

THIS investigation has attempted to explore the stability or variability of the numerous response-determinants of subjects' concepts on the Rorschach Ink Blot Test, to discover the extent to which mere repetition of the test, a change of administrators, or an alteration of the affective tone of the testing situation would bring about alterations in number and content of responses, in determinants, and in location areas.

Related studies are extremely few and their contradictory results are explainable as a function of their decidedly different experimental designs. A single study, that of the Army Air Forces, is closely related to the present experiment in its analysis of variation in number of responses as a function of examiner-differences.

What other response categories are subject to fluctuation as a function of examiner-differences? If examiners studiously attempt to structure three different affectively loaded testing situations, will there be consistent differences, or any differences, as a function of these deliberately varied rapport situations? Will mere repetition of the Rorschach significantly alter any of the scoring categories, regardless of administrator or of affective loading of the administration? If so, which scoring categories are sensitive to such variation?

Thirty-six male college sophomores between the ages of nineteen and twenty-seven participated in a series of three Rorschach tests, administered without testing of the limits. Each test was administered by a different female, and each testing situation was affectively differently loaded. One test was given in the standard manner without pre-tests and without any attempt to give the subject a marked feeling of acceptance or rejection. Another test was administered following two pre-tests during which, by manner and tone of voice, the subject was made to feel rejected and a failure. Still another test was administered after a similar pre-test situation during which the subject was made to feel accepted and successful. Each examiner gave twelve tests first, second and third; twelve tests positively loaded, negatively loaded, and neutral without affective loading. Each subject's different test was administered by a different examiner. An equal number (twelve) of negative, positive, and neutral tests was given on the first, second, and third administrations and by the three different examiners. This plan netted thirty-six Rorschach protocols on each of the three administrations, thirty-six by each of the different examiners, and thirty-six under each of the three different affective loadings of the testing situation. The one hundred eight Rorschach protocols were scored and separately analyzed according to the number of the administration, the administrator, and the type of administration.

On the basis of the analyses of the one hundred eight Rorschach records, the following findings are reported: Subjects who were coldly received tended to misunderstand, and make errors in following directions on simple card-sorting tests. Subjects warmly greeted did not misunderstand, request repetition of, or make errors in following simple directions.

The Experience Balance, or M : sumC ratio, was unstable, only 30% of the subjects maintaining a constant balance throughout the variations in administra-

tion. The remainder shifted emphasis from one side of the ratio to the other at least once.

Ten, thirty-one, and forty-eight *t*-ratios were found at the 1 per cent, 5 per cent, and 10 per cent levels of confidence respectively. Of these forty-eight measurable differences, eight were a function of repetition of the test, thirteen a function of the variation in affective loading of the test situation, and twenty-seven a function of examiner differences.

Of the twenty-three Rorschach response-categories subjected to statistical analysis, three were relatively unaffected by the extra-test, or "situational" stimuli: Dd + S%, Fk, and A. The other twenty categories showed frequency differences at least once at the 1 per cent, 5 per cent, or 10 per cent levels of confidence under at least one of the experimentally varied situations.

In conclusion, performance on the Rorschach, as measured by the frequency of responses in the Rorschach categories, varies significantly with repetition of the test, with variation in the negative or positive rapport conditions of the administration, and with examiner differences. In the present study, the largest and most frequent variations in Rorschach performance were associated with examiner differences.

APPENDIX

EXPLANATION OF SCORING SYMBOLS USED IN THIS STUDY*

Location

W Whole blot
W, S Whole blot and white space used (tabulated as main W and additional S)
D Large usual detail
D, S White space used in addition to D (tabulated as main D and additional S)
d Small usual detail
Dd Unusual detail (or unusual combinations of usual areas)
S White space

Content

H Human figures
Hd Parts of human figures, not anatomical
A Animal figures
Ad Parts of living animals

Popularity

P Popular responses

Determinants

R Response
M Figures in human-like action
FM Animals in animal-like action
m Abstract or inaminate movement
k Shading as three dimensional expanse projected on a two dimensional plane
K Shading as diffusion
FK Shading as three dimensional expanse in vista or perspective
F Form only, not enlivened
Fc Shading as surface appearance or texture, differentiated
C′ Achromatic surface color
FC Definite form with bright color
CF Bright color with indefinite form
C Color only

$$\text{sumC} = \frac{\text{FC} + 2\text{CF} + 3\text{C}}{2}$$

* After Klopfer, Bruno and Helen H. Davidson, *The Rorschach Method of Personality Diagnosis*. Individual Record Blank (New York: World Book Company, 1942).

REFERENCES

1. BECK, S. J. *Rorschach's test: I. Basic processes.* New York: Grune & Stratton, 1944.
2. BECK, S. J. *Rorschach's test: II. A variety of personality pictures.* New York: Grune & Stratton, 1945.
3. BELLAK, L. The concept of projection. *Psychiatry*, 1944, **7**, 365-366.
4. BERGMAN, M. S., GRAHAM, H., & LEAVITT, H. C. Rorschach exploration of consecutive hypnotic chronological age level regressions. *Psychosom. Med.*, 1947, **9**, 20-28.
5. BOCHNER, RUTH, & HALPERN, FLORENCE. *The clinical application of the Rorschach test.* New York: Grune & Stratton, 1945.
6. BRUSSEL, J. A., & HITCH, K. S. The Rorschach method and its uses in military psychiatry. *Psychiat. Quart.*, 1942, **16**, 3-27.
7. BUHLER, CHARLOTTE, BUHLER, K., & LEFEVER, D. W. *Development of the basic Rorschach score.* Los Angeles: mimeographed ed., 1948.
8. BUHLER, CHARLOTTE, & LEFEVER, D. W. A Rorschach study on the psychological characteristics of alcoholics. *Quart. J. Stud. Alcohol*, 1947, **8**, 197-261.
9. COFER, C. N. Psychological test performance under hyoscine: A case of post-infectious encephalopathy. *J. gen. Psychol.*, 1947, **36**, 221-228.
10. DOUGLAS, ANNA GERTRUDE. A tachistoscopic study of the order of emergence in the process of perception. *Psychol. Monogr.*, 1947, **61**, No. 6.
11. FOSBERG, I. A. How do subjects attempt fake results on the Rorschach test? *Rorschach Res. Exch.*, 1943, **7**, 119-121.
12. FOSBERG, I. A. Rorschach reactions under varied instructions. *Rorschach Res. Exch.*, 1938, **3**, 12-31.
13. HALL, CALVIS S. Diagnosing personality by the analysis of dreams. *J. abnorm. soc. Psychol.*, 1947, **42**, 68-79.
14. HEIDER, F. Social perception and phenomenal causality. *Psychol. Rev.*, 1944, **51**, 358-374.
15. HERTZMAN, M., & PEARCE, JANE. The personal meaning of the human figure in the Rorschach. *Psychiatry*, 1947, **10**, 413-422.
16. HSÜ, E. H. The Rorschach responses and factor analysis. *J. gen. Psychol.*, 1947, **37**, 129-138.
17. HUNT, W. A. The future of diagnostic testing in clinical psychology. *J. clin. Psychol.*, 1946, **2**, 311-317.
18. JACOB, Z. Some suggestions on the use of content symbolism. *Rorschach Res. Exch.*, 1944, **8**, 40-41.
19. KADINSKY, D. Human whole and detail responses in the Rorschach test. *Rorschach Res. Exch.*, 1946, **10**, 140-144.
20. KAMMAN, G. R. The Rorschach as a therapeutic agent. *Amer. J. Orthopsychiat.*, 1944, **14**, 21-27.
21. KLOPFER, B., & KELLEY, D. M. *The Rorschach technique.* New York: World Book Co., 1946.
22. LEVINE, K. N., GRASSI, J. R., & GERSON, M. J. Hypnotically induced mood changes in the verbal and graphic Rorschach: A case study. *Rorschach Res. Exch.*, 1943, **7**, 130-144.
23. LEVINE, K. N., GRASSI, J. R., & GERSON, M. J. Hypnotically induced mood changes in the verbal and graphic Rorschach. Part II: The response records. *Rorschach Res. Exch.*, 1944, **8**, 104-124.
24. LINDNER, R. M. Content analysis in Rorschach work. *Rorschach Res. Exch.*, 1946, **10**, 121-129.
25. LINDNER, R. M. Some significant Rorschach responses. *J. crim. Psychopathol.*, 1944, **5**, 775-778.
26. LINDQUIST, E. F. Statistical analysis in educational research. Boston: Houghton Mifflin, 1940.
27. LUCHINS, A. S. Situational and attitudinal influences on Rorschach responses. *Amer. J. Psychiat.*, 1947, **103**, 780-784.
28. PIOTROWSKI, Z. A comparative table of the main Rorschach symbols. *Psychiat. Quart.*, 1942, **16**, 28-37.
29. RORSCHACH, H. *Psychodiagnostics. A diagnostic test based on perception.* Berne, Switzerland: Hans Huber, 1942. English ed. by Paul Lemkan and Bernard Kronenberg.
30. SCHACHTEL, E. G. Subjective definitions of the Rorschach test situations and their effect on test performance. *Psychiatry*, 1945, **8**, 419-448.
31. SCHACHTEL, E. G. Review, *The Rorschach technique*, by Bruno Klopfer. *Psychiatry*, 1942, **5**, 604-606.
32. U. S. ARMY AIR FORCES AVIATION PSYCHOLOGY PROGRAM RESEARCH REPORT, J. P. GUILFORD, editor. *Printed classification tests.* Report No. 5. Washington, D. C.: Government Printing Office, 1947. Chapter 24, Clinical type procedures.
33. WILKINS, W. L., & ADAMS, A. J. The use of the Rorschach test under hypnosis and under sodium amytal in military psychiatry. *J. gen. Psychol.*, 1947, **36**, 131-138.